COMPUTER-ASSISTED SLEEP STAGING
BASED ON SEGMENTATION AND CLUSTERING

Rajeev Agarwal[12*] *and Jean Gotman*[13]

[1]*Stellate Systems, Montreal Canada*
[2]*Concordia University, Montreal, Canada*
[3]*Montreal Neurological Institute, McGill University, Montreal, Canada*

ABSTRACT

In this paper, we present a method that can be used to automatically classify sleep states in an all-night polysomnogram (PSG) to generate a hypnogram for the assesment of sleep-related disorders. The method is based on ideas of segmentation and classification (clustering) using sleep related features. Segments are clustered to generate groups of similar patterns that can subsequently be labeled as one of the accepted clinically relevant sleep stages. Each PSG is processed independently to generate classes of similar patterns in an unsupervised manner, thus achieving pseudo-natural classes that are independent of any classification criterion. Overall performance as compared to manual scoring of 12 subject is shown to be 61.1%.

1. INTRODUCTION

Multi-modal electrographic measurements that include electroencephalogram (EEG), electromyogram (EMG), electrooculograms (EOG) along with other signal types are known as polysomnograms (PSGs). Such measurements are used in the diagnosis and treatment of sleep-related complaints. Normal healthy sleep is organized into four basic biological states (Awake, Light Sleep, Deep Sleep and Rapid-Eye Movement (REM) Sleep) that typically cycle every 60-90 minutes. Depending on the criterion adopted, different stage classification can be derived from these biological states. Current practice in most laboratories is to use the stage classification described by Rechtschaffen and Kales (RK) 1968 [1]. The PSG is generally divided into epochs of 10, 20, 30 or 60 second (referred to as staging epochs), which are then visually classified into one of the RK stages. The resulting temporal evolution of epochs in terms of stage classification is termed the Hypnogram.

A basic premise to sleep staging the PSG is that a pattern is assumed to exist for a time interval until a new pattern emerges signaling the change of state. Since the state of sleep is a continuum from light sleep to deep sleep, the artificial demarcation of sleep stages by the RK classification is a simplification. This characterization of sleep is a methodological concept that attempts to standardize analysis across reviewers and laboratories. Due to the continuum of sleep, the exact time of change from one stage to another is highly subjective and leaves significant room for interpretation by the scorer. Not surprisingly the same scorer will score the transitional epochs differently on dif-

ferent occasions [2]. Many studies have shown inter-scorer agreement ranging from 67-91% [3, 4]. Visual scoring of two healthy subjects in 10 laboratories in Japan showed 67% to 75.3% agreement [5]. Most data on inter-scorer agreement are based on normal subjects.

In addition to being subjective, the visual scoring is very tedious and time consuming. To deal with this, in the last 30 years, several investigators have addressed the idea of computer classification of PSGs, see for example [2, 3, 6]. Although some of these studies show acceptable performance, they are limited to select populations and none have found a common place in clinical settings. Much of this can be attributed to the fact that stage classification schemes have proven to be too ambiguous to be translated into mathematical models. It may also be related to the fact that some of the proposed approaches require threshold and algorithm adjustments for different patient groups.

In this paper, we present an automatic sleep staging method that generates pseudo-natural stages, which can be subsequently classified according to the RK or any other staging classification criterion preferred by the operator. Since user input is required in classifying each of the natural stages, we have termed this approach Computer Assisted Sleep Staging (CASS).

2. METHOD

Subjects

The presented method was developed using 12 all night sleep recordings of subject group that included both males (9) and females (3) with differnt sleep related complaints (8 normals and 4 with different pathologies). The age of the subjects ranged form 17 to 63 years. For each recording, the manual staging was done at least 2-3 years prior to the start of this study.

Computer-Assisted Sleep Staging Method

Unlike some of the methods previously presented in the literature, our approach analyses the complete PSG and only then epochs are classified into one of the valid sleep stages. As suggested by RK criterion, five channels of the PSG are used, namely: two EEG channels (one central derivation and one occipital derivation), a sub-mental EMG channel and two EOG channels (left and right). The stage classification depends on the properties of signals in the complete recording. The method relies on the principles of segmentation and self-organization [7] to cluster the different patterns present in the PSG. Clusters of segments (each cluster represents a particular pattern type) can then be assigned a predefined stage. This latter step allows the interactive participation of the user to customize the staging to their preferences.

---

*Corresponding author. Current address: Stellate Systems, 345 Victoria ave., Suite 300, Montreal, Canada, H3Z 2N2 Tel: (514)486-1306 Loc. 118, Fax: (486-0694); email: ragarwal@stellate.com

Report Documentation Page

Report Date	Report Type	Dates Covered (from... to)
25 Oct 2001	N/A	-

Title and Subtitle	Contract Number
Computer-Assisted Sleep Staging Based on Segmentation and Clustering	Grant Number
	Program Element Number

Author(s)	Project Number
	Task Number
	Work Unit Number

Performing Organization Name(s) and Address(es)	Performing Organization Report Number
Stellate Systems Montreal Canada	

Sponsoring/Monitoring Agency Name(s) and Address(es)	Sponsor/Monitor's Acronym(s)
US Army Research, Development & Standardization Group (UK) PSC 802 Box 15 FPO AE 09499-1500	Sponsor/Monitor's Report Number(s)

Distribution/Availability Statement
Approved for public release, distribution unlimited

Supplementary Notes
Papers from 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, October 25-28, 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.

Abstract

Subject Terms

Report Classification	Classification of this page
unclassified	unclassified

Classification of Abstract	Limitation of Abstract
unclassified	UU

Number of Pages
4

The overall method can be split into two phases. Phase A is the preprocessing component where the PSG is decomposed into variable duration stationary multi-channel segments from which various sleep related features are evaluated. In phase B, the variable duration segments are organized into homogeneous clusters (based on the primitive sleep-related features) that are subsequently assigned a sleep stage to generate the preliminary hypnogram.

Phase A: Segmentation

The five channels of the PSG data are simultaneously broken-down into quasi-stationary segments. The stationarity of the segments allows a meaningful estimation of the features in the next step. Care must be taken to ensure that the segments are relatively unchanging while long enough to allow meaningful estimates of features. With this in mind we have set the minimum duration to be 3 seconds. It was suggested in [8] that frequency-weighed energy, as evaluated by the Nonlinear Energy Operator (NLEO), is sensitive to changes in amplitude and frequency of a signal. We use this idea to simultaneously segment [9] all five channels of data. The actual segmentation criterion is generated using the two EEG and the EMG data in the multi-channel extension of the adaptive segmentation approach in [9] and the segmentation boundaries so generated are extended to include the EOG channels. Segmentation criterion is generated for each of the three signals as

$$G_{nleo}^i(n) = \left| \sum_{m=n-N+1}^{n} \Psi(m) - \sum_{m=n+1}^{n+N} \Psi(m) \right| \quad (1)$$

where $\{i = 1, 2, 3\}$ representing the three data channels and $\Psi(n) = x^2(n) - x(n-1)x(n+1)$, the output of the NLEO. Final segmentation criterion is based on

$$G_{nleo}(n) = \sum_{i=1}^{3} G_{nleo}^i(n) \quad (2)$$

Peak detection is applied to obtain the segment boundaries shown in Figure 1.

Phase A: Feature Extraction

Features describing sleep-related attributes are evaluated for each segment of the PSG. In addition to these features, we also evaluate the maximum absolute amplitude in each segment for each type of signals (EEG, EMG and EOG) and these data are used in an artifact rejection scheme.

Through experimental observations, we selected the following features to represent each PSG segment: amplitude measure, dominant rhythm measure and frequency-weighted energy (FWE) for each of the two EEG channels and the EMG channel, presence of spindles in the central EEG channel, Alpha-Slow-Wave Index (ASI) for the occipital EEG channel, Theta-Slow-Wave Index (TSI) for the central EEG channel and the presence of eye-movement (EM) in the EOG channels. Each segment is, therefore, parameterized by a 13-dimensional feature vector. These features were experimentally determined to provide a good separation of different sleep patterns. The following describes each of these features: Amplitude is defined as the average of the absolute value of the signal. Dominant Rhythm is determined as the pole frequency of the estimated second order autoregressive (AR) model. FWE is defined as the expected value of the output of the NLEO. Presence of spindles is assessed using the ratio of power in the sigma band (11.5-15

Hz) to total power, relative to the background EEG. ASI is the ratio of power in the alpha band (8-11 Hz) to the combined power in delta band (0.5-3.5 Hz) and theta band (3.5-8.0 Hz) . TSI is the ratio of power in the theta band to the combined power in delta and alpha bands. Presence of EMs is determined by the detection of phase reversal of sufficient amplitude in the left and right lowpass filtered (5 Hz cutoff) EOGs.

Phase B: Clustering

The multi-channel PSG segments are clustered into groups with similar properties using the feature vectors generated in Phase A. To do this, an *ad hoc* iterative self-organization scheme [7] based on the *k-means* clustering algorithm is used. A fundamental problem with *k-means* algorithm is the lack of knowledge of the number of pattern types (clusters) present in the data. This is particularly problematic with data sets where no clear separation exists between pattern types. PSG data is one such data type where the state of sleep is a continuum from light to deep sleep. In such data types any number of homogeneous clusters of segments can be created by the self-organization, however the clinical significance may be questionable. The correct number is dictated by the staging standard as well as the PSG in question (e.g., a patient may not have SWS or REM Sleep). An important outcome of this procedure is that no matter how many clusters are generated, they are always relative to the PSG in question. The newly created clusters are not referenced to a particular standard, thus leaving the clinician to apply whatever clinical significance he chooses. In addition to the above problems, self-organization methods are heavily plagued by starting seeds and outlier data members.

The idea in our approach is to start with a number of clusters greater than the number of presumed clusters in the data, and then to reduce the number of clusters by merging those close to each other until the desired K clusters are obtained. The final number of clusters can be selected depending on the subject and the data. In keeping with the RK standard, a good choice is 8, as it accounts for the six key RK sleep stages and two for redundancies (e.g., two types of stage 2 sleep or cluster splitting).

To enhance performance of self-organization, the data are preprocessed by removing artifact-contaminated segments prior to clustering. The artifact (or outlier) segment removal strategy is a multi-layered one that is applied at several stages of the clustering scheme.

Feature Conditioning

Prior to clustering the presence of EM is transformed into 2 types: EMs in the presence of high alpha activity (ASI greater than unity) and those in the presence of low alpha activity (ASI less than unity). The argument being that alpha activity in Wake state is higher than in REM sleep.

Segment features used for clustering are a mixture of different types and have different scales. Prior to clustering, it is therefore necessary that all the features are scaled such that the weighting of any one feature does not play a more important role than any other. Each feature is normalized with respect to the maximum feature value.

Artifact Rejection

The first level of artifact rejection is applied prior to clustering. All segments in which the maximum absolute amplitude of any of the three signal types exceeds a preset threshold are considered to be artifact contaminated and rejected from further consideration. In our prototype, we have used

300, 325 and 200 μVs for the EOG, EMG and EEG data types, respectively. In addition to the amplitude threshold, a technique based on the distribution of the FWE of the EMG and EEG signals is used to identify artifact contaminated segments (see Agarwal *et al.* [7] for details). This is based on the assumption that the FWE of the valid EEG is concentrated in some subspace of the FWE feature space. Segments contaminated by artifact contain higher frequencies and amplitudes and are on the periphery or beyond the EEG (and EMG) subspace of FWE. Since most of the segments throughout the night do not contain artifact, they will form a denser concentration than the artifact contaminated segments. This fact is used to separate segments in which the EEG or EMG is artifact contaminated, thereby achieving artifact rejection.

Phase B: Stage Assignment

Once the segments have been clustered, the user preference of stage classification is incorporated into the algorithm by labeling each cluster with a sleep-stage. This is accomplished in two steps. First, one of the K cluster labels is assigned to each staging epoch depending on which type of pattern (defined by the segment clusters) occupies the largest fraction of the epoch. This procedure imparts segment cluster information onto the staging epochs and generates clusters of staging epochs. By grouping epochs with same cluster labels, we have essentially translated the cluster of segments based on primitive features to cluster of epochs based on these same features. Second, it is necessary to assign a clinically relevant stage label (according to RK or any other classification) to each cluster of staging epochs thus generating the preliminary hypnogram. This is done by selecting five (or fewer) representative epochs from each cluster and asking the operator to score them. The results are used to score all remaining staging epochs within each cluster. By this process, we have accommodated the operator's preference in staging.

3. RESULTS/DISCUSSION

The performance was assesed on an epoch by epoch basis. In its current state the auto staging does not differentiate between stages 3 and 4. Therefore, in the performance assessment, stages 3 and 4 of the manual scoring were combined to yield a single Slow-Wave Sleep (SWS) state. Overall agreement between automatic staging and manual staging was 61.1 % for the 12 subjects.

Figure 2 shows the results of auto scoring at various stages of the method. Figure 2a shows the temporal evolution of segments in each cluster. Note that segments within each cluster are generally confined to cohesive blocks of time. The transition from one cluster to another appears to be relatively sharp indicating possible changes of state. The temporal overlap of some cluster types indicates two possibilites. First, the clusters that overlap may represent similar sleep state (i.e. a sleep state has been split into two or more clusters). Due to the limited graphical resolution this will appear as multiple pattern types within the same time period. Second, two patterns types may vascillate within a block of time period. An example of this is REM period where subperiods may look very much like stage 2 sleep without spindles. Figure 2b shows the temporal evolution of staging epochs in terms of the cluster number to which they belong. This is similar to Figure 2a except that it shows the temporal evolution of different PSG activities in terms of staging epochs. The dashed vertical lines indicate the transition of the epochs between clusters. The segments have been grouped to form staging epochs. Each epoch clusters has been labeled (left of the figure) as one of the valid RK stages. Note Movement stage is not used. Note also that 4 of the 8 clusters are labeled stage 2, suggesting either there exist variants of stage 2 or there is cluster splitting. Figure 2c shows the resulting auto generated hypnogram by merging clusters with same stage labels. Figure 2d shows the manually staged hypnogram. An epoch-by-epoch comparison of the computer generated and the manually staged hypnogram yields a 75.4% agreement. More importantly, the profile of the computer generated hynogram appears to be very similar to the manually staged.

REFERENCES

[1] A. Rechtschaffen and A. Kales, eds., *A manual of standardized terminology, techniques and scoring system for sleep stages of human subjects*, (Los Angeles: Brain Information Service/Brain Research Institute, University of California), 1968.

[2] N. Schaltenbrand *et al.*, "Sleep stage scoring using the neural network model: Comparison between visual and automatic analysis in normal subjects and patients," *Sleep*, vol. 19, pp. 26–35, 1996.

[3] J. M. Gaillard and R. Tissot, "Principles of automatic analysis of sleep records with a hybrid system," *Computers and biomedica research*, vol. 6, pp. 1–13, 1973.

[4] S. Kubicki, L. Holler, I. Berg, C. Pastelak-Price, and R. Dorow, "Sleep EEG evaluation: A comparison of results obtained by visual scoring and automatic analysis with the oxford sleep stager," *Sleep*, vol. 12, pp. 140–149, 1989.

[5] Y. Kim, M. Kurachi, M. Horita, K. Matsuura, and Y. Kamikawa, "Agreement in visual scoring of sleep stages among laboratories in Japan," *J Sleep Res*, pp. 58–60, 1992.

[6] L. E. Larsen and D. O. Walter, "On automatic methods of sleep staging by EEG spectra," *Electroencephalography and Clinical Neurophysiology*, vol. 28, pp. 459–467, 1970.

[7] R. Agarwal, J. Gotman, D. Flanagan, and B. Rosenblatt, "Automatic EEG analysis during long-term monitoring in the ICU," *Electroencephalography and Clinical Neurophysiology*, vol. 107, pp. 44–58, 1998.

[8] J. F. Kaiser, "On a simple algorithm to calculate the 'energy' of a signal," in *Proc. IEEE Int'l Conf. Acoust., Speech, and Signal Processing.*, Apr. 1990.

[9] R. Agarwal and J. Gotman, "Adaptive segmentation of electroencephalographic data using nonlinear energy operator," in *Proceedings IEEE ISCAS '99*, (Orlando, Florida), June 1999.
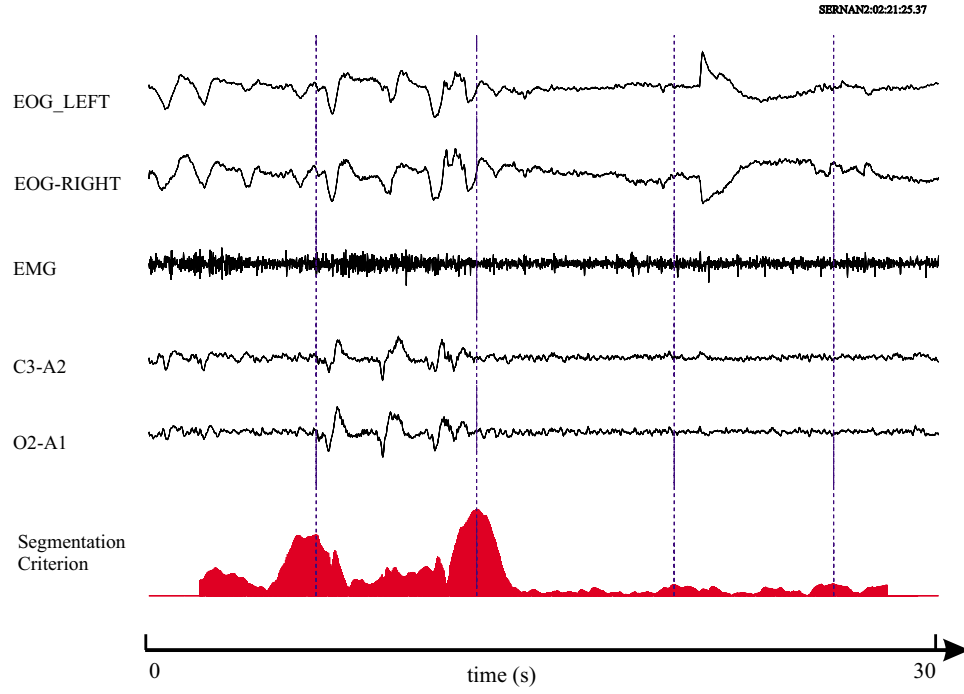
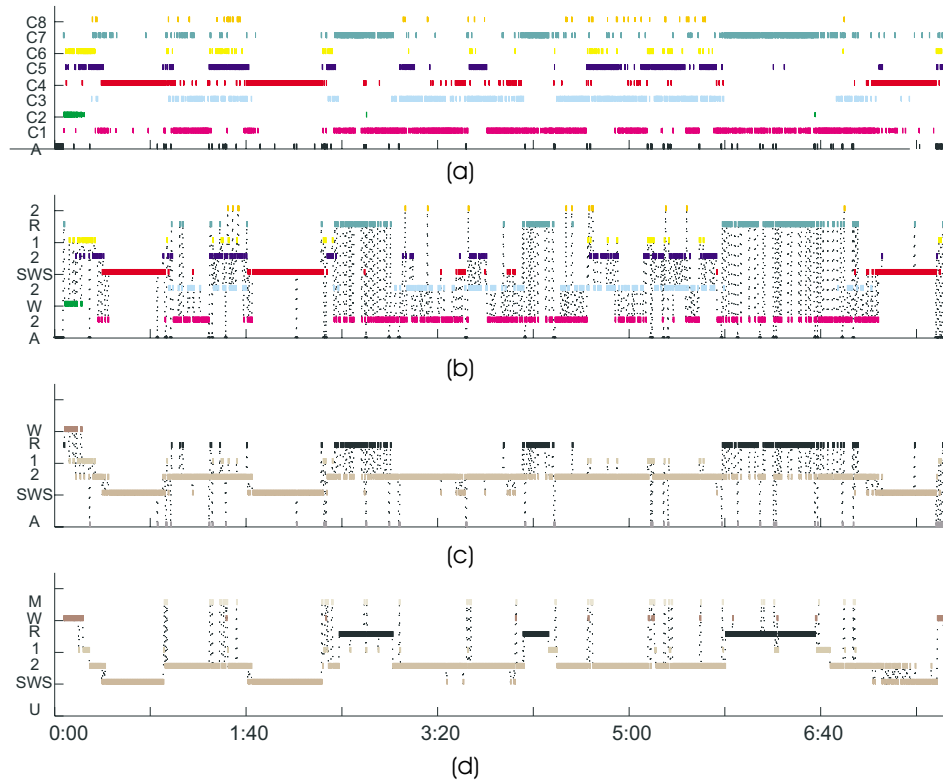**Figure 1. Example of PSG segmentation using the NLEO.**

Figure 2. Generation of hypnogram at various stages of the method. (a)Temporal profile of segment clusters (b) Temporal profile of staging epoch clusters (c) Computer generated hypnogram. (d) Manually staged hypnogram.